

25<sup>th</sup> Meeting of the Wiesbaden Group on Business Registers  
- International Roundtable on Business Survey Frames

Tokyo, 8 – 11 November 2016

*Sean Crick, Luisa Ryan, Jenny Foster and Tracey Rowley*  
*Australian Bureau of Statistics*  
*Session No. 2*

*Role of Business Registers*

***ABS Registers Transformation Program***

## **Introduction**

1. The Australian Bureau of Statistics (ABS) has commenced a five year transformation program that will modernise the infrastructure, systems and processes used to produce official statistics. The program will improve ABS capability and responsiveness, ensure continued delivery of trusted world class statistics and better support decision-making and innovation by providing more timely and integrated statistical services.
2. The ABS Transformation Program provides an opportunity for the role of registers to be expanded beyond their traditional role of providing source information for survey frames to supporting microdata analysis.
3. This paper details the future vision for registers in the ABS, and discusses their role both as a source of information for survey frames, and their expanded role as linking keys to support microdata analysis. An overview of the future conceptual model for the ABS Business Register is included.

## **Overview of ABS Registers**

4. The ABS currently has two statistical registers: the ABS Business Register and the ABS Address Register.
5. The ABS Business Register is used within the ABS as the definitive register of businesses operating in Australia. Historically, its purpose has been to provide a source for the production of accurate and consistent survey frames for business-related statistical collections conducted by the ABS.
6. The scope of the ABS Business Register is all legal entities undertaking productive activity in Australia's economic territory. The information on the ABS Business Register is primarily sourced from the Australian Business Register (maintained by the Australian Tax Office), other data kept by the Australian Tax Office and via ABS profiling of large and/or complex businesses. The ABS uses an Economic Units Model to describe the structure of businesses on the ABS Business. More information on the Economic Units Model can be found in Appendix 1.

7. A common frame is produced from the information extracted from the ABS Business Register each quarter. Survey frames are extracted from the common frame and supplied to various areas of the ABS to support the cycles of different business survey collections.

8. The ABS began development of an Address Register in 2011, with its first use being to provide the frame for the 2016 Census of Population and Housing. The scope of the ABS Address Register is all physical locations with an address in Australia, complemented by other information including geospatial information (geocodes), and land usage/dwelling structure information (residential, commercial, agricultural etc).

9. In 2016, plans were endorsed by the ABS to support the move away from area-based sampling for household surveys towards direct selection from an address register.

10. Currently, accessing the full breadth of micro data held by ABS is difficult, with most data held in siloed repositories and managed by different statistical areas. Sharing of data between areas is ad-hoc, and re-use of data held to create statistical products is rare. To help address these challenges, the ABS is extending the role of registers, expanding them to function as centralised ‘spines’ for microdata analysis via the linking of datasets across both the economic and social domains, in addition to their traditional role as the source of survey sample frame information.

11. The ABS registers of the future will be quite different from the standalone and siloed systems in place today. Registers in the ABS transformed environment will not be held in a single structured database, rather, they will be created virtually, using data linking. Statistical registers will in essence be an environment that encompasses core or ‘spine’ data, along with the capabilities to link to various datasets. These linkable datasets will include datasets that can link back to the spine, the attributes and activities of statistical units, the reasoners (rules engine) that enable derivations and transformations for further linking, plus the ability for views and outputs to be created. The supporting environment will be sufficiently flexible to enable new datasets to be related to the spine as they are created or obtained.

#### **Role of registers as a frame source**

12. Registers will continue to be used in the production of survey frames. The ABS Business Register is well established in this role, with a quarterly common frame being produced from which the majority of economic survey frames are drawn. The ABS Address Register to this point in time has only been used as the frame source for the 2016 Census of Population and Housing; development work is now being undertaken to allow it support the ABS Household Survey Program.

13. The current ABS Business Register focusses on producing units which correspond well to the scope of the Australian Business Register (maintained by the Australian Taxation Office). The future ABS Business Register will also enable the creation of an institutional view of businesses and it will include units which are not recognized by the Australian Business Register as undertaking an enterprise.

14. There are a number of ABS economic surveys which do not currently source their frame from the ABS Business Register. As part of the ABS transformation program, work will be undertaken to migrate or link these remaining collections with the ABS Business Register. The expected benefits of this migration activity are to provide greater coherence between directly collected data, and to provide a platform for microdata analysis.

15. The expanded role of the ABS Business Register, combined with the establishment of an ABS Address Register, allows for improved respondent and provider management. Currently, information on providers selected in ABS collections is held in different systems across different collections and is difficult to extract and analyse. Management of the registers in the transformed environment will enable a more effective co-ordination of information, and will allow for more informed decisions relating to sample selections and improved relationships with data providers and respondents.

16. One of the key benefits registers can provide in the future is supporting the coherence of information relating to similar populations. Registers by their central integrating role will also make data that is already available in the ABS more visible and accessible to users.

#### **Role of registers as linking keys supporting microdata analysis**

17. Internationally, there is growing acknowledgement that registers should become the backbone of economic and social statistics, in particular through better enabling microeconomic analysis, data integration and geospatial analysis. In Australia, accessing and analysing information at the micro level is becoming increasingly important to meet client and policy needs, particularly to inform small area, geospatial and longitudinal analysis. In the transformed environment registers will sit at the heart of the ABS microdata environment, providing the linking keys and foundational unit information.

18. Geospatial data is critical to understanding firm level performance and economic activity, as well as person mobility and access to services. Regional and small area data is in high demand, yet the ABS is not able to meet all of these information needs with directly collected information. Data integration enabled through the transformation of registers provides a potential solution to this, by bringing together information from multiple datasets for the location of interest from both survey and administrative sources.

19. Understanding people's characteristics and attributes and firm level performance over time has become increasingly important for policy evaluation. Longitudinal information has been collected by the ABS in the past on different topics, predominantly by selecting panels and collecting information on a regular basis. Examples of this are the Longitudinal Study of Australian Children and the panel information to create the Business Longitudinal Database Confidentialised Unit Record File – more information on these datasets can be found on the ABS website.

20. To effectively analyse longitudinal data, further information will be needed to help understand changes over time. For the ABS Business Register, this includes improving the data available to help develop a better understanding of the relationship between units, and changes to business structures

over time and the reporting of associated data. For the ABS Address Register this includes improving the data available to help develop an understanding of the relationship between types and use of dwellings and the interrelationship between data sources. Both registers will require detailed maintenance strategies to ensure the ongoing quality of the information included.

21. One of the key features of the transformed corporate infrastructure is a move to metadata driven business processes. This change in focus will place concepts, standards and business rules at the forefront of statistical processes. Register linking keys will provide core infrastructure that will facilitate the reuse of existing information held in data stores, identification of linkages across datasets, increased coherence, and underpin the organisational shift to analysing and interpreting statistical data.

22. The simplification in storing register information provides the opportunity to investigate recent advances in data linking such as semantic web technology. A pilot study was undertaken using the ABS Business Register, with investigations focussing on the feasibility of using semantic methods and technologies to create and maintain a statistical register of business entities. The pilot indicated that while it could be technically possible to create a register using these technologies, the semantic web techniques are not sufficiently understood or advanced within the ABS to provide an effective alternative at this point in time. Semantic techniques do, however, have potential to support broader data linking initiatives and data exploration, which the ABS will continue to explore.

23. The longer term vision for registers is for an external view to be available, allowing users access and capability to link their unit record data with approved ABS Business Register data. The extent to which this vision can be achieved will be considered further taking into account the ABS' commitments to ensuring privacy and confidentiality.

#### **Conceptual model underpinning the ABS Business Register**

24. This section outlines the conceptual model that will underpin the ABS Business Register in the future.

25. A simplified, small foundational dataset will be created as the core of the ABS Business Register. This is referred to as the integrating spine. The integrating spine will contain:

- a comprehensive and exhaustive list of businesses (organisations or entities engaged in economic activity) in scope of the ABS Business Register; and
- identifiers required to link information together. These identifiers are sometimes referred to as keys and have been selected to facilitate unit level linkages to a wide range of statistical and administrative based economic unit record data.

26. The decision was made to keep the spine to a minimal set of data items to avoid additional maintenance activities associated with ensuring the coherence of information across datasets; and to reduce the complexity of the ABS Business Register model. The simplicity of the model is expected to

future-proof the ABS Business Register with flexibility to accommodate future change, whether it is internal or external to the ABS.

27. The spine will include a primary linkage file and two secondary linkage files:

- the primary linkage file will contain Australian Commonwealth Government identifiers of legal entities operating in Australia. An individual unit may have some or all of these identifiers.
- the statistical unit linkage file will contain the information relevant to the ABS economic units model. It will also include a synthesised institutional sector unit to support the institutional view of the economy, and facilitate mappings between institutional and production units.
- the geospatial linkage file will contain a range of geographical identifiers to facilitate the linkage of data based on geography via the Address Register.

28. The datasets (e.g. tax data, profiling data) that are integral to the maintenance of the ABS Business Register, for frame creation and to support economic survey frame operations, will be linked to the spine as the information is required. An overview of this conceptual model is included in Appendix 2.

29. Outputs from the ABS Business Register will be created through linking the relevant datasets with the integrating spine; then extracting the relevant information. Examples of outputs include:

- views/interrogations of the statistical register (to support survey operations, further data linking or register maintenance processes)
- Common and survey frames
- Counts of businesses
- Linkage keys which will be used to link further datasets and create/enhanced views.

30. The base statistical unit of the ABS Business Register will continue to be the legal entity. The legal entity statistical unit is very similar to the System of National Accounts institutional unit. A detailed definition of the legal entity can be found in Appendix 1.

## **Implementation**

31. The current implementation plan for the registers scheduled migration to the new corporate infrastructure is early 2018, with the expectation that they will be available to support business survey frames for the 2018-19 financial year.

## **Conclusion**

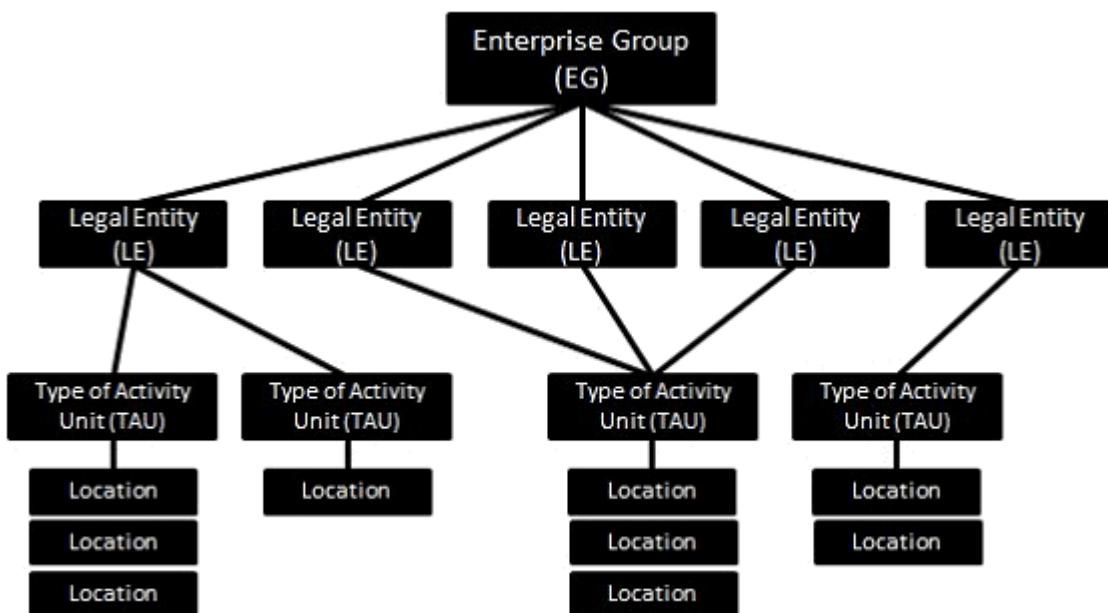
32. The ABS' Transformation Program provides a unique and exciting opportunity to review the role registers play in official statistics and to integrate them into other parts of the statistical cycle. The introduction of linking keys to support microdata analysis will go a long way to maximising the utility of the data held, through the creation of new statistical and analytical sources, and improving coherence of statistics.

33. Being able to link social and economic data in the future through the register integrating spines will enable more advanced geographical and longitudinal analysis, including microdata simulations and analysis at a finer level of detail, for example at the person or firm level.

34. The concept of registers supporting both frame creation and microdata analysis in the future will broaden the horizon for the production of official statistics using richer datasets in order to better inform policy debates.

## Appendix 1 – ABS Economic Units Model

The ABS Economic Units Model that is used to determine the structure of businesses is consistent with Australia's Corporations Law and with the definition of institutional units outlined in 2008 System of National Accounts (SNA). The model consists of: the Enterprise Group (EG), one or more Legal Entities (LE), one or more Type of Activity units (TAU), and one or more locations.



The **Legal Entity (LE)** statistical unit is defined as a unit covering all the operations in Australia of an entity which possesses some or all of the rights and obligations of individual persons or corporations, or which behaves as such in respect of those matters of concern for economic statistics. Examples of legal entities include companies, partnerships, trusts, sole (business) proprietorships, government departments and statutory authorities. Legal entities are institutional units.

The **Enterprise Group (EG)** is an institutional unit covering all the operations within Australia's economic territory of legal entities under common control. Control is defined in Corporations legislation. Majority ownership is not required for control to be exercised.

The **Type of Activity Unit (TAU)** is a producing unit comprising one or more legal entities, sub-entities or branches of a legal entity that can report productive and employment activities via a minimum set of data items.

A **Location** is a single, unbroken physical area, occupied by an organisation, at which or from which, the organisation is engaged in productive activity on a relatively permanent basis, or at which the organisation is undertaking capital expenditure with the intention of commencing productive activity on a relatively permanent basis at some time in the future.



## Appendix 2 – ABS Business Register Spine Concept

